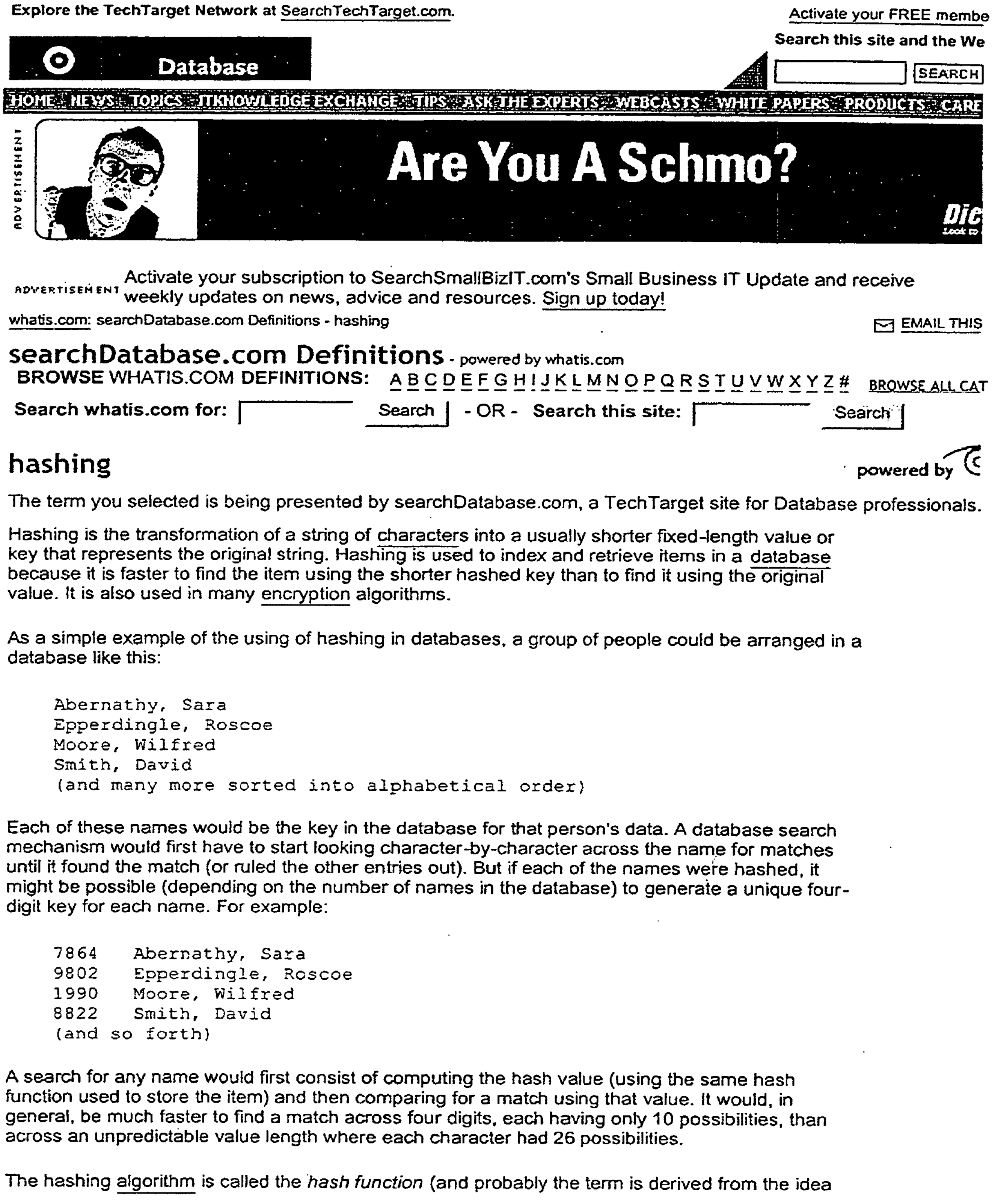
APPENDIX A



Database
SEARCH
HOME NEWS TOPICS ITKNOWLEDGE EXCHANGE TIPS ASK THE EXPERTS WEBCASTS WHITE PAPERS PRODUCTS CARE
ADVERTISEMENT
Are You A Schmo?
Dic

whatis.com: searchDatabase.com Definitions - hashing

✉ EMAIL THIS

## searchDatabase.com Definitions - powered by whatis.com

BROWSE WHATIS.COM DEFINITIONS: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z # BROWSE ALL CAT

Search whatis.com for: [ ] Search - OR - Search this site: [ ] Search

## hashing

powered by

The term you selected is being presented by searchDatabase.com, a TechTarget site for Database professionals.

Hashing is the transformation of a string of characters into a usually shorter fixed-length value or key that represents the original string. Hashing is used to index and retrieve items in a database because it is faster to find the item using the shorter hashed key than to find it using the original value. It is also used in many encryption algorithms.

As a simple example of the using of hashing in databases, a group of people could be arranged in a database like this:

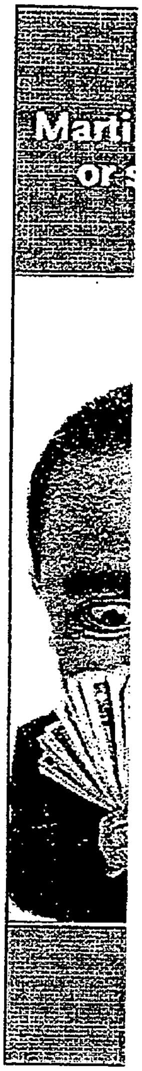
```
Abernathy, Sara
Epperdingle, Roscoe
Moore, Wilfred
Smith, David
(and many more sorted into alphabetical order)
```

Each of these names would be the key in the database for that person's data. A database search mechanism would first have to start looking character-by-character across the name for matches until it found the match (or ruled the other entries out). But if each of the names were hashed, it might be possible (depending on the number of names in the database) to generate a unique four-digit key for each name. For example:

```
7864   Abernathy, Sara
9802   Epperdingle, Roscoe
1990   Moore, Wilfred
8822   Smith, David
(and so forth)
```

A search for any name would first consist of computing the hash value (using the same hash function used to store the item) and then comparing for a match using that value. It would, in general, be much faster to find a match across four digits, each having only 10 possibilities, than across an unpredictable value length where each character had 26 possibilities.

The hashing algorithm is called the *hash function* (and probably the term is derived from the idea that the resulting hash value can be thought of as a "mixed up" version of the represented value). In addition to faster data retrieval, hashing is also used to encrypt and decrypt digital signatures (used to authenticate message senders and receivers). The digital signature is transformed with the hash function and then both the hashed value (known as a message-digest) and the signature are sent in separate transmissions to the receiver. Using the same hash function as the sender, the receiver derives a message-digest from the signature and compares it with the message-digest it also received. They should be the same.



The hash function is used to index the original value or key and then used later each time the data associated with the value or key is to be retrieved. Thus, hashing is always a one-way operation. There's no need to "reverse engineer" the hash function by analyzing the hashed values. In fact, the ideal hash function can't be derived by such analysis. A good hash function also should not produce the same hash value from two different inputs. If it does, this is known as a *collision*. A hash function that offers an extremely low risk of collision may be considered acceptable.

Here are some relatively simple hash functions that have been used:

- The division-remainder method: The size of the number of items in the table is estimated. That number is then used as a divisor into each original value or key to extract a quotient and a remainder. The remainder is the hashed value. (Since this method is liable to produce a number of collisions, any search mechanism would have to be able to recognize a collision and offer an alternate search mechanism.)
- Folding: This method divides the original value (digits in this case) into several parts, adds the parts together, and then uses the last four digits (or some other arbitrary number of digits that will work ) as the hashed value or key.
- Radix transformation: Where the value or key is digital, the number base (or radix) can be changed resulting in a different sequence of digits. (For example, a decimal numbered key could be transformed into a hexadecimal numbered key.) High-order digits could be discarded to fit a hash value of uniform length.
- Digit rearrangement: This is simply taking part of the original value or key such as digits in positions 3 through 6, reversing their order, and then using that sequence of digits as the hash value or key.

A hash function that works well for database storage and retrieval might not work as for cryptographic or error-checking purposes. There are several well-known hash functions used in cryptography. These include the message-digest hash functions MD2, MD4, and MD5, used for hashing digital signatures into a shorter value called a message-digest, and the Secure Hash Algorithm (SHA), a standard algorithm, that makes a larger (60-bit) message digest and is similar to MD4.

>> Find products and vendors related to hashing.

**Read more about it:**

>> What Is a Hash Function? , an explanation from RSA Security, mentions the MD2, MD4, MD5, and SHA algorithms.
>> This lecture on Hashing requires some math background, but may help you understand what hashing is.

**Last updated on:** Mar 12, 2002

« Back to previous page Go to whatis.com home page »

**WHAT'S NEW**
on searchDatabase

1. SQL Server basics learning guide
2. Visit the Expert Answer Center
3. Free database white papers
4. Register now for free newsletters!

HOME NEWS TOPICS ITKNOWLEDGE EXCHANGE TIPS ASK THE EXPERTS WEBCASTS WHITE PAPERS PRODUCTS CARE

About Us | Contact Us | For Advertisers | For Business Partners | Reprints SEARCH

SearchDatabase.com is part of the TechTarget network of industry-specific IT Web sites

**WINDOWS**
SearchExchange.com
SearchVB.com
SearchWin2000.com
SearchWindowsSecurity.com
Labmice.net
MyITForum.com

**APPLICATIONS**
SearchCRM.com
SearchSAP.com

**ENTERPRISE IT MANAGEMENT**
SearchCIO.com
SearchSmallBizIT.com

**CORE TECHNOLOGIES**
SearchDatabase.com
SearchMobileComputing.com
SearchNetworking.com
SearchOracle.com
SearchSecurity.com
SearchStorage.com
SearchWebServices.com
WhatIs.com

**PLATFORMS**
Search390.com
Search400.com
SearchDomino.com
SearchEnterpriseLinux.com

TechTarget Expert Answer Center | TechTarget Enterprise IT Conferences | TechTarget Corporate Web Site | Media Kit

Explore SearchTechTarget.com, the guide to the TechTarget network of industry-specific IT Web sites.

Read our P